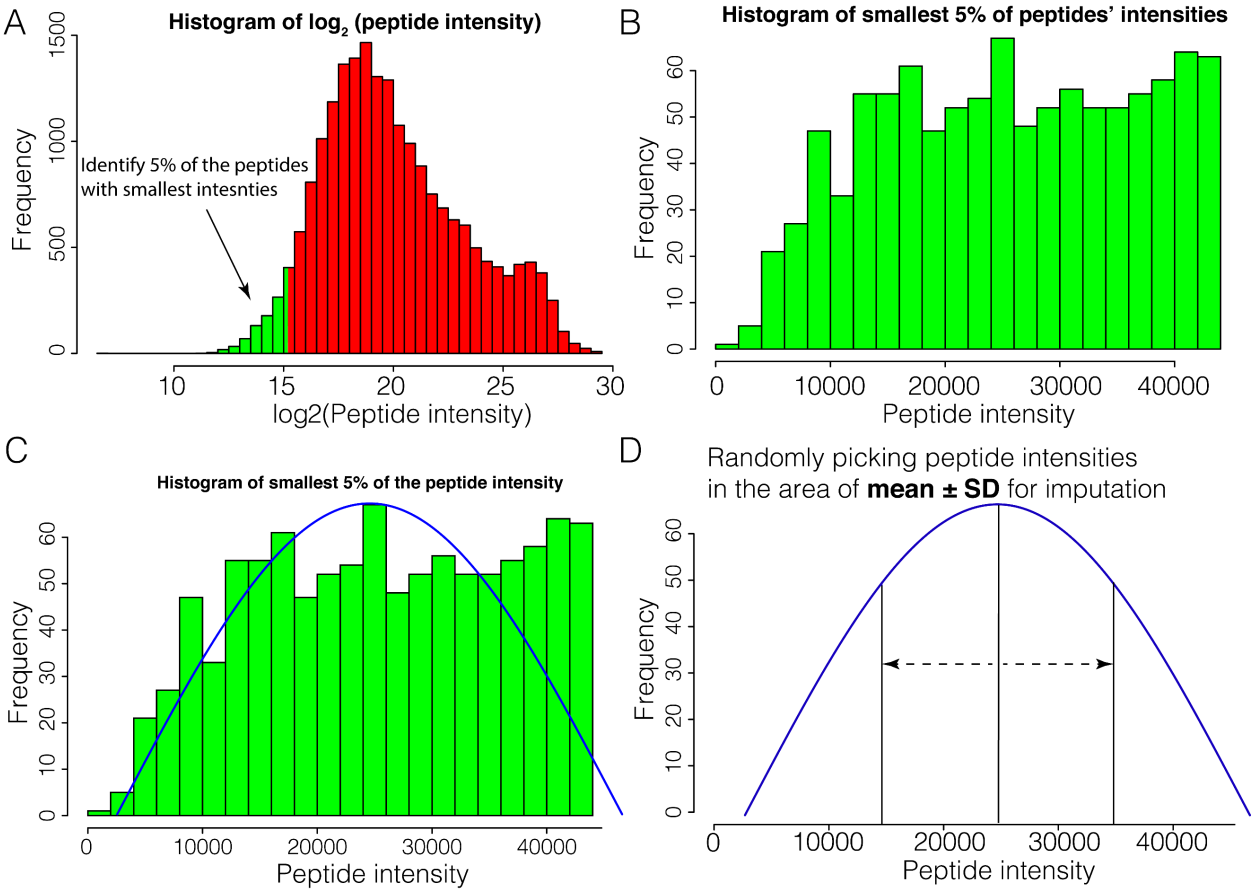
We hypothesized that peptides that are in high abundance are consistently detected in replicate assays. Peptides that are very low in abundance are less consistently detected in replicate assays due to them being close to or below the detection limit of the instrument. We assume the detection limit of the instrument can be represented by the smallest 5% of the data. To impute these missing or zero values with a non-zero value we performed the following steps.



1. Using data from 15 experiments (3 technical replicates each with 5 timepoints), we ranked all detected peptides from smallest to largest using their intensity values. There was a total of 20,491 peptides whose intensity values were not missing or not zero in this study. We identified 1025 peptides that correspond to the smallest 5% category (Panel A, green bars). Panel B shows a histogram of the distribution for the smallest 5% of non-zero intensity values.
2. We generated a normal distribution based on the smallest 5% data (Panel C). Note that no fitting is done at this step. The distribution function is generated by simply calculating the mean (μ = 25660.32) and SD (σ = 11093.32) of these data.
3. Missing values were then imputed with a peptide intensity value that was randomly picked within the area of mean ± sd (Panel D). In this way, the smallest numbers within this 0-42,000 range were less frequently picked as they will skew the calculations of enzyme kinetics, and largest numbers were less frequently picked as they should be detected if peptides were at this level of abundance.